September 14, 2012

CDF Long Term Data Preservation in Italy

S. Amerio[1], L. Dell'Agnello[2], G.Punzi[3], L.Ristori[4], D. Salomoni[2], F.Scuri[4]

[1] *INFN Padova* [2] *CNAF* [3] *Scuola Normale Superiore di Pisa* [4] *INFN Pisa*

Abstract

We present a project for the long term preservation of CDF data in Italy. The goal of the project is to store a complete copy of CDF data at CNAF computing center and to provide the necessary services for data access and analysis in the long term future (more than 10 years from now). The project - the first involving INFN and CNAF - develops in parallel to the CDF data preservation plan being designed at Fermilab and complements it, providing an additional copy of data offsite and access resources. After a general description of the current status of CDF project at Fermilab, we will describe in detail the INFN-CNAF data preservation plan, providing an estimate of manpower and costs and a tentative time schedule.

Contents

1 Introduction **2**

2 Tevatron data: motivations for long term preservation **3**

3 Goal of CDF data preservation project **5**

4 Status of CDF data preservation project **5**

5 CDF data at CNAF: motivations **7**

6 Description of the project **8**

6.1 Phase 1: Data copy from Fermilab to CNAF 8

6.1.1 Current status of CDF data at Fermilab 8

6.1.2 Data to be copied at CNAF 9

6.1.3 Technical implementation 10

6.2 Phase 2: Maintenance of data access and related services 11

6.2.1 Current services at CNAF 12

6.2.2 Future 12

7 Resources **13**

7.1 Requests for CNAF storage and computing resources 13

7.2 Manpower requests for data transfer, data maintenance, and user support 13

8 Cost estimate **14**
8.1 Option 1: all data copied in 2013 14
8.2 Option 2: data copy splitted in two years 15
8.2.1 2013 15
8.2.2 2014 15

9 Time schedule **15**

10 Appendix A: Raw data to be copied **16**

11 References **16**

1 Introduction

Data collected in High Energy Physics (HEP) experiments are the result of a significant human and financial effort. The preservation of HEP data beyond the lifetime of the experiment is of crucial importance for several reasons:

- long term completion and extension of scientific programs: it is demonstrated (LEP experiments, BaBar) that about 5 to 10% of total scientific production of a collaboration is produced after the end of data taking; often these analysis are more sophisticated and can exploit the physics potential of the entire collected dataset;

- the possibility of performing cross collaboration analysis, analyzing data from several experiments at once; in this way statistical and/or systematic uncertainties of single experiments can be reduced, and new analyses performed. Such an effort to combine analyses is already ongoing, for example between the H1 and ZEUS collaborations, and an evaluation of such an approach is underway between the Belle and BaBar collaborations.

- Perform new analysis on data from past experiments with new theoretical models or new analysis techniques; this can lead a significant increase in precision for the determination of physical observables. As an example data from JADE experiment at PETRA e^+e^- collider have been recently re-analyzed leading to several precise measurements proving the running of the strong coupling constant in a unique energy range [1].

- Education, training and outreach: past HEP experiment data can be analyzed by graduate and undergraduates students from institutions different from the ones collaborating to the experiment; this is a unique opportunity to reach a wide audience, and an invaluable tool for particle physics classes.

Since 2008 a study group on Data Preservation in High Energy Physics (DPHEP) has been constituted to investigate the technical and organizational aspects of HEP data preservation. DPHEP identifies different models of data preservation, of increasing complexity [2], from the simpler preservation of the capability of providing additional documentation on published analysis (extra-data tables, internal notes, etc.), to the more complex preservation of the reconstruction and simulation software and basic level data. The latter is recommended, as the only way to provide the full physics analysis chain and retain full flexibility for future use.

The experiments BaBar, Belle, BES-III, CLAS, CLEO, CDF, D0, H1 and ZEUS are represented in DPHEP, the LHC experiments ALICE, ATLAS, CMS and LHCb having also recently joined. The associated computing centers at CERN (Switzerland/France), DESY (Germany), Fermilab (USA), IHEP (China), JLAB (USA), KEK (Japan) and SLAC (USA) are all also represented in DPHEP.

Since Spring 2012 a dedicated task force has been created at CDF to study in detail the current CDF computing model and to develop a proposal for long term preservation of data and analysis capabilities. INFN is one of the main participants of CDF experiment and has always supported its computing needs, with dedicated CPU and storage resources at CNAF computing center in Bologna. In this document we present the proposal for the participation of INFN in the CDF data preservation project: the idea is to host at CNAF a complete copy of CDF data and all the related services (access portal, analysis software, ...); this will protect against accidental loss or damage of data files, and will allow INFN will to have an indipendent copy of CDF data in Italy.

2 Tevatron data: motivations for long term preservation

CDF and D0 experiments at the Tevatron collected 10 fb^{-1} of data during RunII. These data are still yielding valuable information on the nature of physics phenomena and will continue in the future, especially in light of new discoveries by LHC or other experiments and advances in theoretical models. Additionally, as no plans are foreseen for a proton-antiproton collider in the future, Tevatron data will be unique for a very long time in terms of initial state particles and of measurements of effects enhanced by $q\bar{q}$ interactions. In these areas, the Tevatron will remain competitive with the LHC.

As examples within the realm of top physics, $t\bar{t}$ production asymmetry measurements have shown a discrepancy with the Standard Model which could hint at new physics [3]. At the LHC such effects are more difficult to observe, as symmetric gg-initiated events dominate top pair production.
Moreover, differing production mechanisms in the two environments test distinct aspects of the Standard Model and require different analysis strategies, as for example in $t\bar{t}$ spin correlation measurements [4].
Tevatron data will also remain of importance for single-top searches [5], particularly in the s-channel which is more challenging at the LHC than at Tevatron, because the relative cross section is much smaller.

The mass of the top quark is now known with a relative precision of 0.54%, limited by the systematic uncertainties, which are dominated by the jet energy scale uncertainty [6]. This is the result of the combination of several measurements made by CDF and D0 in different $t\bar{t}$ decay channels on data samples with integrated luminosity up to 5.8 fb^{-1}. Uncertainty on jet energy scale is expected to improve as analysis are performed on the full data samples, since analysis techniques constrain the jet energy scale using kinematical information from $W \to qq'$ decays. For the first time the total uncertainty of the combination is below 1 GeV; such a level of precision urges to study theoretically in more details the exact renormalization scheme definition corresponding to the current top mass measurements. We have entered the era of precision measurements in the top sector, and the Tevatron will provide a substantial contribution to the top mass world average for many years to come.

In the electroweak sector, one of the most important measurements made at the Tevatron is the precise determination of the W mass. In conjunction with top mass, the W boson mass constrains the mass of the Higgs boson, and possibly new particles beyond the standard model. The measurement is very challenging due to presence of an undetected neutrino from the W decay, which makes it impossibly to fully reconstruct the final state unambiguously. Recently CDF and D0 have measured the most precise values of the W mass to date [7], achieving a total uncertainty of 19 and 23 MeV/c^2 respectively, dominated by the uncertainty on parton distribusion functions (PDFs). It will be difficult for the LHC to overtake this precision for at least several years. Moreover, the current measurements are based on exposures of 2.2 and 4.3 fb^{-1} in integrated luminosity at CDF and D0, respectively. With the full data sample, these measurements could constrain systematic uncertainties and, in principle, reach a precision of 10 MeV/c^2. Attaining this precision will require considerable effort, especially in reducing the uncertain on PDFs.

Heavy flavor physics has several potential analyses that can be carried out in the coming years. Some of the ideas that have emerged include measuring A_{SL} in B^0 and B_s^0 decays, studying the forward-backward asymmetry in charm and bottom production, measurements of the interference between scalar and vector resonances in B decays, and measuring production cross sections and polarizations (where possible) for as many heavy flavor states as possible. There are numerous decay modes that can be extracted from the data, some which will likely be overlooked by other experiments.

More generally, Tevatron data might be useful to cross check a result from another experiment. This may be particularly important in the light of new discoveries at the LHC, which may require CDF data to be revisited, possibly with new, more advanced analysis techniques. This was recently demonstrated with the evidence for a CP asymmetry difference between $D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ decays from the LHCb experiment that was soon after confirmed by CDF [8].

It as also to be stressed that Tevatron measurements are made in a unique energy domain, which will be no longer available; therefore QCD measurements performed on Tevatron data will continue to be as valuable in understanding QCD as LHC ones. Examples are measurements on diphoton cross section, $Z/W+$ jets, underlying and minimum bias events, diffractive W and Z production. For the same reason, before the definitive shutdown of the accelerator, data were collected at two different energy points, 900 and 300 GeV. The data samples, collected by minimum-bias and selective triggers, will provide some valuable legacy measurements in non-perturbative QCD, soft and strong interactions.

These are only some highlights of the enormous potential of Tevatron data. Tevatron will keep producing high quality scientific results, though at a lower rate in the coming years. But Tevatron will also serve as a fundamental point of comparison for LHC.

3 Goal of CDF data preservation project

The goal of CDF data preservation project is to preserve CDF data and the complete analysis capability in the long term future, till the CDF Collaboration will continue to exist and beyond. It has also to prepare for the phase when the CDF data set will enter the public domain. In this phase, access to the data should be preserved at least in a simplified format, with special consideration to its potential value for outreach and education.

As far as data are concerned, only one copy of CDF data currently exists and it is stored at Fermilab. Small subsets of data (less than 10%) are stored offsite, at CNAF and KISTI computing centers. To protect against data loss and/or corruption at least one copy of the data has to be stored offsite and proper validation and recovery procedures set up.

All the code to produce analysis-level ntuples from raw data has to be preserved, in case reprocessing is needed. Production of physics results should not be more difficult in the future than it is at present despite the expected diminishing support and reduced availability of expert advice. In case new theories need to be tested on CDF data, the framework for MonteCarlo production of simulated events should be preserved, and new generators should be easily interfaced to the detector simulation. Users will need computing resources to run their jobs, so a long term job submission framework has to be foreseen as well.

4 Status of CDF data preservation project

In Spring 2012 a task force dedicated to data preservation has been created at CDF. It is composed by CDF members from different instutions, joined by computing experts from Fermilab Computing Sector. It is expected to submit for approval by the Collaboration a written report by September 1^{st}, 2012.

Four main areas of work have been identified, and corresponding working groups created: data preservation and access, code preservation, long term job submission and preservation of documentation. For each area, the current status of CDF computing is analysed in detail to highlight the minimum requirements needed for long term preservation; possible solutions for the long term future are being investigated.

A detailed plan has not been approved yet. Preliminary outcomes of CDF data preservation task force studies are:

- data access: it is desirable to keep the current access method (Sequential Access via Metadata, SAM [9]) for data access in the long term future. It is possible to maintain the current version of SAM, developed for RunII Tevatron experiments, or migrate to the future version which is being developed for the Intensity Frontier (IF) experiments at Fermilab. The first option requires less manpower in the short term, but depends on the long term availabilty of external code; the second solution requires instead immediated action by Fermilab SAM experts to adapt CDF code to the new SAM features, but will require minimum maintenance effort in the long term future;

- data preservation: any project of data preservation needs to ensure as much redundancy of data as possible. It would be useful if a copy CDF data was stored also offsite, in one or more computing centers. This will ensure maximum protection against accidental data loss or corruption. The task force is in contact with CNAF and KISTI computing centers to discuss their involvement in CDF data preservation project to host a copy of CDF data. Metadata (run condition, configuration, trigger, luminosity, alignment and calibration information) stored in Oracle databases have to be preserved in case data need to be reprocessed or new MC samples generated;

- code preservation: all CDF code runs on Scientific Linux 5 (SL5) operating system and a new SL6 version will be available in 2013; a possible solution for the long term future is to exploit virtualization techniques and run CDF code on SL5 or SL6 virtual machines. Security issues may arise when SL6 will not be supported anymore (after 2020) but can be bypassed keeping a small subset of machines for CDF needs beyond a firewall. The possibility to exploit grid resources should be investigated.

- job submission: a job submission framework similar to the CDF one is being developed for the IF experiments at Fermilab. A possible solution for the future is to migrate CDF job submission code to the IF framework; CDF users will be allowed to run their jobs on Fermilab computing farm. In the far future, when security issues may prevent CDF jobs from running on Fermilab grid, a set of machines can be dedicated to CDF jobs. Given the rapid increase of computing performances, after 2020 only a small fraction of the current computing resources will be needed to process CDF data.

- documentation: it is necessary to preserve detailed documentation about all the fundamental ingredients of CDF analysis, from raw data processing to analysis results. A new web page has to be created, where information contained in current CDF webpage can be re-organized.

5 CDF data at CNAF: motivations

The preservation of a copy of CDF data at CNAF is important for several reasons, regarding not only CDF collaboration, but INFN as well.

The CDF data set is the result of a big investment of money and human resources over the course of more than 30 years and a significant fraction of these resources were provided by INFN. The INFN investment in CDF between 1980 and 2012 is difficult to quantify precisely:a rough estimate is of the order of 30 million Euros plus 1000 FTE*year.

Given this large contribution of INFN funds and human resources, and given the important role played in CDF, we expect INFN to feel the responsibility to make sure that the product of such a big investment be preserved in the safest possible way and the data be kept available to the italian scientific community for many years to come. We believe that INFN should consider this very seriously as a fundamental institutional responsibility and that it would not be appropriate to leave the decision on if and for how long our data must be preserved exclusively in the hands of Fermilab. We also regard having a mirror archival location in Europe as a sensible and necessary safety measure. We must also realistically consider that remote access to Fermilab computing facilities may not be straightforward in the long term future and may be subject to changing policies of the laboratory or the US government. Even if data access will be assured to CDF researchers as long as the collaboration will exists, access by INFN researchers can be guaranteed in the long run only if data are available under our control within our home institution by having a full copy of all of them archived at an INFN facility. The quality of the accessibility of the data is also an important factor and strongly depends on the hardware and human resources being made available locally, those should therefore be under INFN control to guarantee effective use of the data if needed.

The cost of archiving the CDF dataset is small compared to what was necessary to produce it, and will keep decreasing in the future with the continuing exponential progress of storage technology. We therefore consider the decision to preserve it as extremely sound economically, independent of whether we can prove today that it will actually be needed in the future. The mere possibility that it may be crucial for something we cannot imagine now should be a sufficient motivation.

The preservation of CDF datas is also important in a broader perspective. Preserving CDF data at CNAF computing center does not mean only copying it to tape: it requires to develop a system allowing users in the long term future to access and use that data. The design of such system is a complex project. CDF long term data preservation system at CNAF can serve as a prototype for future experiments which are now supported by INFN and which will soon face the problem of data archival (as

Data group	Volume (TB)
MC (raw data)	1163
MC (ntuples)	624
Data (raw)	1857
Data (production)	3834
Data (ntuples)	1492
TOTAL	8970

Table 1: CDF data volume.

an example, LHC experiments have already started to discuss about data preservation and access in the long term future). It will be an opportunity for CNAF to develop the necessary skills to take a significant role in the long term preservation of data, which is of great interest at international level in high energy physics and beyond.

CNAF has already developed a software (see section 6.2.2) which would allow not-up-to-date code to run on computing resources shared with current experiments, taking into account security issues which may arise. CDF would be a perfect case study to test this solution.

6 Description of the project

6.1 Phase 1: Data copy from Fermilab to CNAF

6.1.1 Current status of CDF data at Fermilab

Three different data formats exist at CDF: *raw data*, *production data*, which have undergone a first reconstruction of physics objects and assigned to a specific dataset depending on the triggers satisfied, and *ntuple-level data*, in three different flavours, for different analysis groups (top physics, B physics and generic). CDF data volume (collected data and Monte Carlo simulated data) is summarized in Table 1 and amounts to about 9 PB. All data are stored on tape (LTO3, LTO4 and T10K) within a dedicated library at Fermilab controlled by the Enstore mass storage system. The data handling system of CDF is based on the Sequential Access via Metadata, SAM [9]. Metadata describing the contents of each data file is stored in the SAM data catalogue. The data handling system can be used to define datasets based upon metadata queries within the catalogue. The files within such datasets can then be delivered upon demand from tape to worker nodes for processing via a 800 TB dCache-based disk cache. dCache [10] fetches files requested by the users and stores them on a distributed pool of disk servers for the user to access over the network. The disk servers are managed as simple cache with the least access file deleted in case space to fetch new files is needed. dCache uses the namespace, PNFS, of enstore and supports multiple access protocols (dcap, kerberized FTP, SRM and Gridftp).

6.1.2 Data to be copied at CNAF

Depending on the available resources, we can envisage different subsets of CDF data that can be copied at CNAF.

1. All raw and ntuple-level data, all production and ntuple-level MC samples (5.1 PB); CNAF will host a complete copy of CDF data and MC samples, completely independent from Fermilab. Ntuples can be immediately accessed and analysed by users, and in case a re-production of files is needed, it can be performed locally. This can be necessary in case a file is corrupted, or in case information is missing. For example, not all jet reconstruction or b-tagging algorithms are run by default on CDF data; often in the past users requested ntuple reproduction to add missing information. In case ntuples need to be reproduced, the re-production would require $\sim$1.3 hours/GB with the present processing architectures.

2. Raw and ntuple-level data only, ntuple-level MC samples (4.0 PB); this solution will allow CNAF to have complete control on CDF data; in case MC ntuples need to be reproduced, CNAF should ask Fermilab for the production files.

3. Raw and ntuple-level data only (3.3 PB); CNAF will have complete control on CDF data. MC ntuples will only be available at Fermilab. As a consequence users can access and analyze data at CNAF but have to rely on Fermilab for MC samples.

4. Raw data only (1.9 PB); CNAF would act as data storage only. Raw data cannot be analysed by users, who have to rely on Fermilab ntuples. In case CNAF wants to produce analysis level files from raw data the following resources will be needed (referred to current processing architectures): $\sim$0.6 hours/GB from raw data to production files, $\sim$1.3 hours/GB to produce ntuples. For MC samples, users can rely on Fermilab only.

Proposal 1 is the most expensive in terms of immediate resources, but it is the only one which guarantees complete independence from Fermilab. Users could perform a complete analysis on datasets stored at CNAF. Proposal 2 can be a good compromise between immediate resources and independence from Fermilab. The lack of MC production samples could be bypassed requesting the samples from Fermilab when needed and re-processing them, or re-generating the MC from scratch. Proposals 3 and 4 are the less expensive in the short term, but the complete absence of MC (and data for solution 4) ntuples at CNAF could be a major difficulty for data analysis in the longer term. It would be necessary to reproduce the MC ntuples in a second step, from scratch or requesting the production samples from Fermilab.

In summary, Proposal 1 is the option we support most, but in case of reduced resources, proposal 2 is an acceptable compromise. Proposals 3 and 4 should be avoided, if possible, as they would require significant resources in the long term future to re-generate the analysis level samples.

Also metadata stored in databases should be copied and preserved at CNAF. The offline database contains run condition information necessary to produce MC events or re-process ntuples from raw data. SAM database contains metadata describing the contents of each data file and it is fundamental to access data. They are Oracle based and amount to about 250 GB.

6.1.3 Technical implementation

The amount of data to be copied varies between 4.0 and 5.1 PB. The copy of such amount data in a reasonable time window ($\sim$ few months) will require the setup of dedicated resources at Fermilab and CNAF.

For the copy, two possible options are being considered:

1. copy driven by CNAF via a local SAM station (*Pull* option). At CNAF a SAM station to access the local copy of CDF data has already been installed. Currently CDF datasets can be copied there using a simple SAM command (*sam_get_dataset*) which fetches the required datasets from Fermilab tape system and copies them on CNAF SAM cache disks via gridFTP protocol. In the current configuration the data copy rate is quite slow (about 1 TB/day) but fast enough for the copy of limited amounts of data. To copy all data in 1 - 5 months, depending on the tape resources, a copy rate of about 1 PB/month (33 TB/day, 380 MB/s) is necessary; this rate can be achieved only through CNAF gridFTP servers but SAM is not designed to perform the copy via third-party servers: SAM code has to be modified to create an interface with CNAF gridFTP servers. This task will be performed by CDF in collaboration with CNAF and SAM experts; we estimate it can be implemented and tested in 1 month, requiring 0.1 FTE on CNAF and CDF side.

   Another issue to be addressed is the interface between CNAF storage system and the SAM station. SAM has been designed to be interfaced with Fermilab tape system, managed by Enstore software. CNAF storage system instead is managed by the Grid Enabled Mass Storage System (GEMSS) software [11]. GEMSS main components are the IBM General Parallel File System (GPFS) [12]; the Tivoli Storage Manager (TSM) [13] software that is used to manage the tape layer access; the StoRM [14] layer that is used in conjunction with the GridFTP servers to provide remote Grid access; and different customized programs for the data migration operation (i.e., data flow from disk to tape), for the data recall process (i.e., data flow from tape to disk) and application data access. It will be necessary to adapt SAM commands to move the data to and from the CNAF tape system. The data flow to the tape system is straightforward: SAM has to copy CDF datasets in an area which is periodically scanned by one of the GEMSS services to create a list of files to be migrated to tape. The migration to tape is taken care by the TSM process. Once files are migrated to tape, SAM catalogue will be updated with the new location of the files. Dataset retrieval instead will be more tricky: *sam_get_dataset* has to be modified to to call the appropriate functions to start the data transfer from tape to disk, in case the files are not stored in SAM cache. Again this task will be performed by CDF in collaboration with CNAF and SAM experts; we estimate it can be implemented and tested in 1 month, requiring 0.1 FTE on CNAF and CDF side.

As far as hardware is concerned, at least two tape drives will be needed at CNAF and two 10 gbit servers to guarantee the copy rate. Moreover, a disk buffer of at least 100 GB will be necessary.

On Fermilab side, this solution does not require any additional setting. Current CDF system (tape drives and dCache) can be exploited, at the cost of a reduction of resources dedicated to CDF users. Tests are planned to provide an exact estimate of the needed resources on Fermilab side, considering different amounts of data to be copied. The exact amount of resources available for the copy will be negotiated with CDF collaboration based on the results of the tests and on the expected data analysis load in 2013-2014.

Data integrity during the copy can be checked by SAM comparing the CRC checksum from the metadata to the one on disk.

2. copy driven by Fermilab via SRM-copy (*Push* option). Fermilab loops over tapes, dumps them to disk and copies them to CNAF. This solution requires to setup a dedicated system at Fermilab. To achieve a copy rate of 380 MB/s, 4 (2) dedicated LTO4 (T10K) drives and 4 file servers are needed. The files can be copied using the *srmcp* command with adler32 checksum end-to-end.

   On CNAF side the same resources as for the Pull options are needed. With respect to the Pull option, it will be necessary to implement an interface between the SAM station and the SRM-client. For file retrieval it will be necessary to adapt SAM commands as for the Pull option. These task will be performed by CDF in collaboration with CNAF and SAM experts; we estimate they can be implemented and tested in 2 months, requiring 0.1 FTE on CNAF and CDF side.

As far as network is concerned, there are two possible options:

- the 2x8.5 Gb/s LHCOPN network between Fermilab and CERN; the usage for CDF data copy has to be negotiated with CMS experiment.
- 10 GE GARR network between CNAF and Fermilab via the StarLight switch in Chicago. This network will be available starting Fall 2012.

6.2 Phase 2: Maintenance of data access and related services

Once data is copied at CNAF, it will be necessary to allow users to access and analyse it. In the following we will review the current status of CDF computing at CNAF and make a proposal for the long term future.

6.2.1 Current services at CNAF

CNAF currently hosts the following CDF computing services:

- SAM station, to copy datasets from Fermilab and access local ones;
- SAM station cache (300 TB), where a subset of CDF data and MC is stored;
- users areas (100 TB);
- job submission portal (Eurogrid) to submit CDF jobs on dedicated resources at Tier1 (8000 HS06) and on the LCG; Eurogrid comprises a headnode and two squid servers (production and backup);
- CDF code volume, accessed via AFS.

All the services are replicas of CDF services at Fermilab, so they required limited manpower for installation. They are installed as virtual machines on SL5 and SL6 operating systems. Two services (SAM station and users areas) are installed on SL4 OS, but the upgrade to SL6 is planned by the end of September 2012. They have stable performance and currently require minimum maintenance (0.3 FTE from CDF and occasional support from CNAF experts).

6.2.2 Future

In the medium term future - until 2015, when CDF is expected to migrate to archival mode - it will be necessary to maintain at CNAF all the current services (SAM station and cache, job submission portal, code and users areas). By the end of September 2012 all services will be upgraded to SL5 or SL6. As SL5 will be supported until 2017 and SL6 until 2020, maintenance in the medium term future will require minimum support.

In the long term future CNAF services will have to migrate to archival mode. The details of CDF long term data preservation model have not been defined yet, but the following services have to be guaranteed to users: data and database preservation and access; access to CDF code and to computing resources for data analysis and MonteCarlo production.

Running CDF legacy code requires addressing several issues, like availability of suitable hardware resources, software maintenance and handling of computer and network security.

CNAF proposes that services used to access CDF data be eventually migrated to a dynamic virtual infrastructure. We plan to implement this infrastructure so that CDF services can be instantiated on-demand on pre-packaged virtual machines (VMs) in a controlled environment, where in- and out-bound access to these services and connection to storage data is administratively controlled. The content of these VMs (operating system, installed programs) will be agreed between CDF and CNAF system administrators. The set-up will be such that, when authorized access to CDF data is requested, instantiation of the virtual services will happen automatically and the

VMs will be placed into a suitably isolated network infrastructure. We propose to realize this dynamic virtual infrastructure through the INFN-developed Worker Nodes on Demand Service (WNoDeS [15]), a framework to implement scalable virtualization infrastructures to be accessed via local, Grid or Cloud interfaces. Besides being a scalable and flexible solution, a big advantage of implementing access to CDF code through WNoDeS is that it will not require provisioning and dedicating hardware resources to CDF, since WNoDeS can be configured to run virtual services using generic CNAF Tier-1 resources.

For what regards the implementation of the WNoDeS-based infrastructure, we propose to start with a pilot phase in early 2013 and continue throughout 2013, trying out access to CDF data copied from Fermilab as it becomes available, and testing secure access to/from the infrastructure. We estimate that, for what regards hardware resources, in the pilot phase we can use existing systems at CNAF. On the personnel side, CNAF will require 0.3 FTE to implement, test and operate the infrastructure.

7 Resources

7.1 Requests for CNAF storage and computing resources

The following storage and computing resources are needed for the implementation of the project at CNAF:

- 4.0 PB (data raw and ntuples, MC ntuples only), 5.1 PB (data and MC, raw and nuples);
- disk, to be used as cache for the copy: 100 TB;
- two T10K drives and two servers dedicated to the copy, to guarantee the maximum copy rate (380 MB/s); after the copy, this hardware will be used also by other experiments supported by CNAF. This will be taken into account in section 8.
- one server to store CDF databases;
- Oracle licence for CDF databases;
- Virtualization software (kvm) licences for the job submission framework.

In case the copy is splitted in two years, one T10K server and one disk server will be enough to guarantee the necessay copy rate.

7.2 Manpower requests for data transfer, data maintenance, and user support

The project will be implemented in collaboration with Fermilab Computing Sector and CDF collaboration. On CNAF side manpower will be required for the following tasks:

- building SAM interfaces to CNAF tape system and gridFTP servers/SRM-client;
- setup of the resources (tape drives and servers) needed for the copy;
- monitoring of the copy;
- maintenance of the current services;
- development and test of a new framework for job submission.

We estimate that 0.5 FTE for 2 years will be required. The CNAF personnel will be supported by a member of CDF Italian collaboration. They will work in strict contact with Fermilab CS experts and CDF data preservation task force.

8 Cost estimate

In the following we will assume the following cost for the data copy:

- tape: 40 euro/TB;
- Usage of tape drive and copy server: 2 euro/TB[1].

8.1 Option 1: all data copied in 2013

- tape: 168 keuro (data raw and ntuples, MC ntuples only), 215 keuro (data and MC, raw and ntuples);
- disk: we don't need additional disk, we can exploit the cache currently used for the SAM station (300 TB + 167 TB in 2013);
- server to copy CDF database: 3 keuro.
- Oracle licence: already covered by CNAF.
- Virtualization software (kvm) licences: already covered by CNAF.
- 2-years post-doc (Assegno di Ricerca): 50 keuros

TOTAL = 221 keuro (data raw and ntuples, MC ntuples only), 268 keuro (data and MC, raw and ntuples).

[1]With two tape drives and two disk servers we can copy 900 TB/month (1 PB/month with an efficiency of 90%); considering a lifetime of 4 years for the hardware, the cost per TB is $\frac{50\,keuro}{48\,months \times 900\,TB/month} = 1.2\,euro/TB$. To be conservative, a cost of 2 euro/TB has been considered throughout this document.

8.2 Option 2: data copy splitted in two years

8.2.1 2013

- tape: 89 keuro (data and MC ntuples, 1492 + 624 TB);
- disk: we don't need additional disk, we can exploit the cache currently used for the SAM station (400 TB + 167 TB in 2013);
- 1-year post-doc (Assegno di Ricerca): 25 keuros
- Virtualization software (kvm) licences: already covered by CNAF.

TOTAL = 114 keuro.

8.2.2 2014

- tape: 76 keuro (raw data, 1857 TB) or 126 keuro (raw data and MC);
- disk: we don't need additional disk, we can exploit the cache currently used for the SAM station (300 TB + 167 TB in 2013);
- 1-year post-doc (Assegno di Ricerca): 25 keuros
- server to copy CDF database: 3 keuro.
- Oracle licence: already covered by CNAF.

TOTAL = 104 keuro (only raw data) or 154 keuro (raw data and MC).

9 Time schedule

- The CDF data preservation project is expected to be ready by the end of 2012. The implementation should start at the beginning of 2013.
- The results of data transfer tests between Fermilab and CNAF are expected to be ready by Fall 2012.
- The data copy to CNAF can be started as soon as the tape and the necessary hardware will be available (Spring 2013). In case it is staged in two years, data and MC ntuples can be copied in Spring 2013, data raw (and eventually MC) can be copied in Spring 2014.
- The development and test at CNAF of the new framework for job submission will start in Spring 2013.
- The migration of all services to archival mode will start not earlier than Spring 2014.

input dataset (Xphysr)	size (TB)	description
aphysr	80	copies of a few high-pt triggers+J/Psi - not used anymore
bphysr	359	leptons
cphysr	207	photons
dphysr	171	events for monitoring, usually not processed
ephysr	438	b-jets, met, tau , zbb
gphysr	264	jets and minimum bias
hphysr	325	B-datasets
iphysr	18	i for inclusive, used for special runs and usually not processed
jphysr	260	B-datasets
TOTAL	2126	

Table 2: CDF raw and production datasets

10 Appendix A: Raw data to be copied

The physics data collected by the CDF trigger system is divided into the the raw data streams listed in Tab. 2, which are then split by the production job into production datasets.

The total raw data amout to 2126 TB. Removing the dataset that are not used for analysis we are left with 1857 TB.

11 References

References

[1] S. Bethke et al., *Determination of the strong coupling S from hadronic event shapes and NNLO QCD predictions using JADE data*, Eur. Phys. J. C 64 (2009) 351 [arXiv:0810.1389].

[2] *Data Preservation in High-Energy Physics*, DPHEP-2009-001 - November 30, 2009, arXiv:0912.0255v1 [hep-ex]

[3] The CDF Collaboration, Conference note 10807, http://www-cdf.fnal.gov/physics/new/top/2012/LepJet_AFB_Winter2012/CDF10807.pdf; V. M. Abazov et al. (D0 Collaboration), Phys. Rev. Lett. 84, 112005 (2011); arXiv:1107.4995.

[4] The CDF Collaboration, Conference note 10719, http://www-cdf.fnal.gov/physics/new/top/2011//SpinCorrDIL/SpinCorrDIL_Pub/spincorrPubnote.pdf; V. M. Abazov et al. (D0 Collaboration), arXiv:1110.4194 [hep-ex].

[5] The CDF Collaboration, Conference Note 10793, http://www-cdf.fnal.gov/physics/new/top/confNotes/cdf10793_SingleTop_7.5_public.pdf; V. M. Abazov et al. (D0 Collaboration), Phys. Rev. D 84, 112001 (2011).

[6] Tevatron Electroweak Working Group, CDF, D0 Collaborations, arXiv:1107.5255 [hep-ex].

[7] R.C.Lopes de Sa (on behalf of CDF and D0 collaborations), arXiv:1204.3260v2 [hep-ex].

[8] The CDF Collaboration, Conference Note 10784, http://www-cdf.fnal.gov/physics/new/bottom/120216.blessed-CPVcharm10fb/cdf10784.pdf

[9] http://projects.fnal.gov/samgrid/WhatisSAM.html

[10] http://www.dcache.org/

[11] D. Andreotti et al., INFN-CNAF Tier-1 Storage and Data Management Systems for the LHC Experiments, 2011 J. Phys.: Conf. Ser. 331 052005

[12] http://www-03.ibm.com/systems/software/gpfs/

[13] http://www-01.ibm.com/software/tivoli/products/storage-mgr/

[14] R. Zappi, E. Ronchieri, A. Forti, and A. Ghiselli, An efficient Grid data access with StoRM, S.C. Lin and E. Yen (eds.), Data Driven e-Science: Use Cases and Successful Applications of Distributed Computing Infrastructures (ISGC 2010), Springer Science + Business Media, LLC 2011

[15] http://web.infn.it/wnodes/index.php/wnodes